

# Why is there explanatory asymmetry?<sup>i</sup>

Johannes Persson, Lund University

## 0. INTRODUCTION

According to the deductive-nomological (D-N) model of explanation, an explanation of a fact consists in its entailment from laws and one or several initial conditions. It fits the following kind of examples nicely. Why did the balloon's volume decrease? Explanation: *When it was submerged in tempered water the pressure increased; and under conditions of constant temperature,  $PV=k$ .* Among the counter-examples that led to the D-N model's rejection, the ones building on its inability to account for explanatory asymmetry were especially effective. In the example, the increase in pressure explains the decrease in volume, but the reverse inference from decrease in volume to increase in pressure would not, in this case, be an acceptable explanation.<sup>1</sup> Yet it fits the D-N model as easily as the first one. Such examples from explanatory asymmetry were not only important in this negative sense, they also had the function of promoting an alternative view, the causal model of explanation. The causal model accounts for explanatory asymmetry by deriving it from the causal asymmetry between cause and effect: The causal asymmetry between submerging the balloon in tempered water which increased the pressure (cause) and the decrease in volume (effect) contributes to the explanatory asymmetry between them. The overall aim of this paper is to examine the claim that

---

<sup>i</sup> Previous drafts of the second part of this paper were read at the Copenhagen backwards causation colloquium 1999, and a symposium of the Nordic network on the theory of explanation, Lund 2000. For valuable comments thanks to Alexander Bird, Phil Dowe, Jan Faye, Bengt Hansson, Henrik Hällsten, Hugh Mellor, Rebecca Schweder, Petri Ylikoski, and the 'small seminar' in Lund. I owe a special debt to Matti Sintonen. Congratulations, Matti! Financial support for writing this paper was provided by Riksbankens Jubileumsfond.

explanation is asymmetrical because causation is asymmetrical. The link between causal and explanatory asymmetry is focussed on. It is argued that many theories of causation account for causal asymmetry in a way that stops a causal model from contributing to our understanding of explanatory asymmetry. What appears to be generally advantageous with causal approaches is normally true only of a few specific causal accounts. These remaining alternatives, however, may well have some less attractive features that would never be brought to mind in the familiar explanatory contexts of balloons, flag poles, and the like—and with a standard concept of causation.

#### 1. WHAT'S THE POINT?

What makes a model of explanation satisfactory? If it has the right consequences, surely. But unless the implausible<sup>2</sup> assumption that a model of explanation has to cover most explanations is added, judging whether one candidate is more satisfactory than another cannot easily be done, as above, by mechanical use of pro- and con-examples. The rejection of inadequate alternatives is less straightforward than the textbook has it, and it is equally difficult to say when a model receives genuine support.

I think that other kinds of tests should be preferred. What could fruitfully be examined is whether the fact that a certain explanation has the characteristics the model is built around *contributes* to its explanatory power. And a minimum requirement for shedding light on that issue is to examine whether these characteristics really exist in and are special to that model.

Why examine this by comparing the superseded D-N model and causal approaches? Hempel thought of causal explanation as a special case of D-N explanation,<sup>3</sup> but in contemporary discussion it has correctly been noticed that causal approaches have to diverge more radically from the D-N approach to resolve the traditional problems. The problem of asymmetry provides a good example. It couldn't be resolved simply by claiming that despite our ways of

expressing the laws, these are really asymmetric. The reason is that sometimes the law is put to work in reverse—as for instance when we explain the increased pressure in cylinders by decreased volume. In order to do better than the D-N model, a causal model of explanation shouldn't be a slight modification of its predecessor but supply asymmetry in its own characteristic way.

However, it can by no means be concluded in advance that causal approaches live up to this goal, i.e. that they have something valuable to offer. Perhaps Hempel was the more realistic about the way causation actually functions in explanatory contexts, and then the comparison between the D-N model and causal approaches might be a particularly good illustration of the difficulties involved in judging when a model of explanation is more satisfactory than its competitors. My paper partly justifies this more sceptic view: Only few causal accounts seem to provide the right kind of asymmetry. More precisely, it seems that only certain kinds of mechanistic views do well enough, and that many traditional views do not perform at all well.

## 2. A CONTEMPORARY ILLUSTRATION

According to Peter Lipton in his *Inference to the best explanation*, the D-N model is in want of two characteristics a causal model has—locality and asymmetry:

Ordinary explanations do not have to meet the requirements of that [deductive-nomological] model, because one need not give a law to give a cause, and one need not know a law to have good reason to believe that a cause is a cause. As for the over-permissiveness of the deductive-nomological model, the reason recession explains red shift but not conversely is that causes explain effects and not conversely (Lipton 1991, 33)

It is illustrative to note how problematic Lipton's analysis is. On the surface, what Lipton says about the characteristics of causation seems true. But only as

long as we rely on a pre-theoretic understanding of causation. Asymmetry will be discussed extensively below, but let me first use locality to briefly present the structure of the argument in this paper. Anyone familiar with theories of causation knows that it is not that clear that “one need not give a law to give a cause”. On all but singularist views causation seems to entail laws. True, one need not give a specific law to give a cause, but according to many a law must nevertheless exist. If this doesn’t threaten the locality of causation and hence of explanation, does the D-N model really have a disadvantage here? The D-N model clearly admits of elliptical explanation. We do not have to cite a law, but to have a *complete* explanation it should be filled in. To put it roughly: a) the causal model’s reliance only on a local link is contentious, and b) the locality traditional theories of causation can contribute to explanations is not clearly different from its “counterpart” in the D-N model. Moving from the surface to a deeper level of understanding causation seems to reconcile the differences with regard to locality between the two models.

Furthermore, to say with Lipton that “the notion of causation is indispensable to philosophy, ordinary life, and much of science, and if we wait for a fully adequate analysis of causation before we use it to analyze other things, we may have to wait forever” (Lipton 1991, 33), provides no rationale for our confidence in the virtues of a causal model unless one has independent reason to believe that the relevant characteristics will be preserved through the analytical transformations. As will be seen below, the transformations often change a lot more than we expect.

### 3. CAUSAL AND EXPLANATORY ASYMMETRY: THE PROBLEMS WITH AGENCY AND MANIPULATION ACCOUNTS

I turn now to a more detailed examination, this time of the claim that causation contributes asymmetry to explanations. Like Lipton above we often take it for granted that causation is asymmetric. Perhaps not obviously so on a general

level—anxiety causes failure and failure causes anxiety, and we have many symmetrical relations in the realm of scientific theories—but surely so on a local level. My state of anxiety at  $t_1$  caused that mistake at  $t_2$  which in turn made me even more anxious when I reached  $t_3$ ; and these cases of local causation are certainly asymmetric according to our intuitions. But as soon as the student of explanation looks for a particular, theoretically more well-developed, theory of causation to go with his explanatory project, causal asymmetry also turns out to be a rather problematic property.

One problem with asymmetry in causation is that it can be a property of such different levels, not equally appropriate for a solution to the problem of explanatory asymmetry. For instance, some causal theorists would like to throw asymmetry in, not as a characteristic of the world but rather of us who take an interest in the world. These are *perspectival* suggestions, and they seem to make the path via causation an unnecessary detour for the student of explanation. The justification of subjectivist or pragmatic perspectives is clearly a more straight forward affair in explanatory contexts than in causal ones. So even if this kind of asymmetry exist in some causal models, it springs from a more fundamental source that, if one accepts it, immediately imposes this kind of asymmetry on, for instance, explanations as well. One example is Huw Price's view, that causal asymmetry is a projection of our own temporal asymmetry as agents (Price 1992, Price 1996). Price's idea is that an especially important domain of our world, the micro-physical world, is symmetric. But since we as readily speak as if causation were asymmetric in micro-physics, the asymmetry must have an external source. It is projected onto the world, Price claims, and then he suggests that it is the means-end relation that is projected onto the causal relation. Now, such asymmetry could be projected directly onto the explanatory relation. On perspectival views, there is no reason to assume that it is the fact that a certain element is causal that contributes asymmetry to the explanation. This is not an objection to perspectival views of causation or explanation, it is a critique of the advocate of causal models of explanation who thinks that it is causation that contributes asymmetry to explanations by

being perspectival.

Not every agency or anthropocentric causation theory is perspectival. Theories such as Gasking's and von Wright's agency centred accounts (Gasking 1955, von Wright 1975) are of a different kind. They are *manipulation* views. von Wright summarises his view the following way: "What makes  $p$  a cause-factor relative to the effect factor  $q$  is, I shall maintain, the fact that by *manipulating*  $p$ , i.e. by producing changes in it 'at will' as we say, we could bring about changes in  $q$ " (von Wright 1975, 107). Both Gasking and von Wright have an "experimentalist" concept of causation. It is not intended to be projected onto situations where manipulation does not occur, and it says nothing about the existence of asymmetries in the world. One supposed problem with these manipulation accounts is exactly that the causal domain according to the intuitions of many includes numerous non-manipulated phenomena as well.<sup>4</sup> As I understand von Wright and Gasking, they are well aware of this limitation and don't try to overcome it: "The notion of 'cause' here elucidated is the fundamental or primitive one. It is not the property of scientists; except for those whose work most directly bears on such things as engineering, agriculture or medicine, and who are naturally interested in helping their practical colleagues, scientists hardly ever make use of the notion" (Gasking 1955, 486). The concept of explanation is different. Even if causes are sometimes expelled from scientific discourse, explanations are seldom restricted only to more practical disciplines—engineering, agriculture and medicine. So we might immediately find reason to doubt that the asymmetry in manipulation accounts of causation is suitable for contributing asymmetry to explanations.

But if we don't commit ourselves to the view that all explanation and therefore all explanatory asymmetry has to be accounted for by a single model, we need to examine whether, in experimental contexts, this causal asymmetry is suitable for an account of explanatory asymmetry. In order to evaluate this partial account of explanatory asymmetry, we need to see what it is that accounts for asymmetry in manipulation accounts of causation.

Gasking seems only interested in distinguishing between cases where we have a manipulative technique “for producing B by producing A” from cases where we in the circumstances can infer B from A. In some cases where we also can infer A from B, we cannot produce A by producing B. Expanding on this view, one option to have asymmetry in manipulation views is to have asymmetric laws of manipulation. Even though we can infer the temperature of a piece of iron by observing that it glows as easily as the other way around, this is not true for manipulation:

We speak of making iron glow by making it hot, i.e. by applying to it the general technique for making things hot, namely, putting on a fire, which in this special case also makes it glow. We do not speak of making iron hot by making it glow, for we have no general manipulative technique for making things glow. And we say that the high temperature causes the glowing, not *vice-versa*. (Gasking 1955, 482)

Following this line of thought is illustrative because we are now again closer to the D-N model. Causation has a close kinship to lawlike regularities. Furthermore, we have already seen that postulating asymmetry at the law-level does not work well for scientific laws in explanatory contexts. Given shifting circumstances many (scientific) laws are explanatory in both directions. So if the asymmetry of manipulation laws is what is supposed to constitute causal asymmetry in manipulation accounts, it is not enough for explanatory asymmetry. We have explanatory asymmetry also in situations where the law works in both directions.

To tie asymmetry to the laws or something else in the context where manipulation takes place also have the probable consequence that the asymmetry of manipulation accounts turns out to relate to the asymmetry of causal properties in the world (that we use in order to manipulate this or that) rather than to the direct fact that we manipulate this or that. Of course, the exact nature of this consequence depends on the special case, and would demand a more detailed argumentation, which however would conflict with the need to see the total picture. A rougher sketch will have to do in this paper. For

the present purposes, it is enough to entertain the possibility that manipulation accounts sometimes can account for asymmetry by relying on more metaphysical theories of causation.

In sum: the perspectival and manipulation ideas do not seem to contribute causal asymmetry to explanations in their own right. Either they seem to contribute a perspective more fundamental than the idea of causation (in which case the contribution doesn't come from causation), or they depend on other more fundamental causal asymmetries, to which I turn below. It should also be noted that many versions of these causal accounts operate with a very restricted causal concept, motivated perhaps for matters of causation but doubtful when suggested as a basis for a model of explanation.

#### 4. CAUSE&EFFECT VIEWS

The traditional way of conceiving of causation in its particular instances is in terms of causes and effects. The two extremely influential Humean accounts of causation (Hume 1777/1902, Sect VIII)—the regularity and the counterfactual view—both tend to understand causation exclusively by its causes and effects.

The cause&effect perspective seems *prima facie* to be intrinsically asymmetric. To be sure, we are not talking about a cause/cause perspective, nor of an effect/effect perspective. So far so good; but, again, when we go on to give a causal account from the cause&effect perspective, the causal asymmetry we took for granted when applying the predicates "cause" and "effect" is not guaranteed to be preserved.

Let us take the regularity view first. An important component of it is that A and B are related as cause and effect only if A precedes B. Since the temporal relation is asymmetric, it determines—at least partly—what the nature of causal asymmetry is in this model. In the absence of other asymmetric relationships between cause and effect it results in a purely temporal account of causal asymmetry. That this is problematic as a causal

account is well-known: while we tend to view the question of whether causation sometimes may be a simultaneous or even a backwards-in-time affair as an open question, the regularity account closes the issue in a premature way.<sup>5</sup> There is also a further difficulty for the student of explanation. If temporal asymmetry is a good candidate for explanatory asymmetry, then what ties it to a causal model? This Humean model of causation simply makes it part of the definition that causes precede their effects; a Humean about explanation could as well simply define that explanans should precede the explanandum, and indeed he sometimes does. We have as good or poor reasons in the one case as in the other. It is not a genuine advance to postulate asymmetry in causation, and then give a causal account of explanatory asymmetry, if the reasons for having this kind of asymmetry are the same on both levels. This Humean account of causal asymmetry does not seem to contribute anything to explanatory asymmetry.

The other Humean approach, the counterfactual view, often approaches asymmetry by pointing to asymmetric patterns in the world. For instance, the way the future is depends on the way the present is. If the present were different, the future would be different. But not so in reverse. “Seldom, if ever, can we find a clearly true counterfactual about how the past would be if the present were somehow different” (Lewis 1979/1986, 32). This is the *branching view* of asymmetry. It is not tied exclusively to counterfactual accounts of causation but fits many other cause&effect views as well. The fork asymmetry views of Reichenbach, Hausman, and others share two important characteristics with Lewis’s view: asymmetry depends on more than one instance of causation, and the asymmetry accounted for in this way is causal asymmetry. For the present purposes, we can refer to all of them as branching views.

My quarrel with branching views is neither with the much discussed problems of back-tracking conditionals in everyday situations, nor with the possibility that the micro-physical world has properties which in fact stops such asymmetric patterns (Price 1992). Let us grant that the direction of the

majority of the open forks in a causal structure determines the direction of its causal relations (Reichenbach 1956), and that as a matter of fact on all levels the majority of open forks of all causal structures are pointed in the same direction. Would we then have a good account of explanatory asymmetry? I urge 'no'. It is not easy to see how branching views could throw any genuine light on explanatory asymmetry. In all cases, the direction of an explanation would depend on the direction of the causal structure, so unless one knows or can imagine also the remote causal future and history of the relevant causal structure, the asymmetry cannot be determined, and this poorly mirrors what explanatory asymmetry is taken to be like. Most of the facts about what happened before A and after B and *vice versa* are irrelevant to the explanatory relations between A and B, and yet according to the branching view they are not. Unlike causal asymmetry (according to branching views, that is), explanatory asymmetry is more sensitive to causal complexity than causal origins and remote consequences.

#### 5. CAUSAL LINES

Russell, sceptical about cause&effect views (Russell 1912/1954, Russell 1948/1951, Ch. IX), sometimes conceived of causation in terms of "causal lines". The idea of a causal line comes from thoughts about enduring things, states, events, and processes. As often, Russell links causation to inference:

I call a series of events a "causal line", if given some of them, we can infer something about the others without having to know anything about the environment. [...] When two events belong to one causal line, the earlier may be said to "cause" the latter. In this way laws of the form "A causes B" may preserve a certain validity. (Russell 1948/1951, 333–334)

The causal line conception thus differs from the cause&effect perspective in admitting more entities than the cause into consideration. It is not even said that the events that we employ have to be earlier than the events we infer (even

if Russell reserves the name “cause” for earlier events),<sup>6</sup> so at this point two differences between cause&effect views and causal lines emerge. The causal-line picture is historically or *diachronically* richer than most cause&effect views in that it admits not only preceding but also simultaneous and succeeding events. The causal line perspective also tends to be *horizontally* richer than cause&effect views in that it admits not just of one cause-effect relation but of several. But while richer in these senses, it is still meagre in the sense we are interested in here. It doesn’t provide causal asymmetry—or if it can be claimed that actual inference-relations are enough for asymmetry: it is multi-directional. In that respect it has exactly the same characteristics as the D-N model. It can therefore not live up to Lipton’s promise: to supply asymmetry in a better way than the traditional but superseded theories of explanation.

6. CAUSATION AS PERSISTENCE, CONSERVED QUANTITIES, AND TRANSMISSION  
 Embedded in Russell’s thoughts is also a complementing idea which should be extracted and related to the suggestions of among others Salmon and Phil Dowe. Russell adds to the foregoing discussion by remarking:

Throughout a given causal line, there may be constancy of quality, constancy of structure, or gradual change in either, but not sudden change of any considerable magnitude. I should consider the process from speaker to listener in broadcasting one causal line: here the beginning and end are similar in quality as well as structure, but the intermediate links—sound waves, electromagnetic waves, and physiological processes—have only a resemblance of structure to each other and to the initial and final terms of the series. (Russell 1948/1951, 477)

The characteristics Russell here focus on are quite independent of the ideas of richness and inference. The present characteristics have to do with similarities between different regions of a causal line. Facts at  $t_1$  resemble those at  $t_2$  if these facts belong to the same causal line. Salmon has developed a more detailed account of process causality where similarity in structure is important

and Dowe's version of the idea is also interesting.<sup>7</sup> Part of Dowe's objective is to make less vague the characteristics that are supposed to be similar throughout a causal line. Russell's and Salmon's vague "structure" is replaced by concepts that the scientific theories themselves might give an exact meaning. (Dowe 1992, 213). This is an advantage. But apart from the fact that current scientific theories might be at least partly mistaken, a problem with the suggestion is that conceptual formation in the sciences are frequently driven by causal considerations. Though concealed, the concept of cause might be a component of some of the scientific concepts. This makes the explanatory power of Dowe's account less transparent than is for instance Russell's. In this context, the additional remark that scientific concept formation to a high degree aims for explanatory power should once again make the student of explanation ask himself whether it is the causal or the unanalysed explanatory element that does the job, should he settle for Dowe's or Salmon's view.<sup>8</sup>

The most pressing problem in this context is to supply the right kind of causal asymmetry. Are preservation and conservation at all asymmetric characteristics? Salmon worries about the more "dynamic" connotations of causation when he says that "In fact, I think that the concept of causal *transmission* is a principal part of a satisfactory explanation of the causal structure of the world, but it is not present in Dowe's theory" (Salmon 1998, 20). It was in his 1984 book that Salmon revived a criterion suggested by Hans Reichenbach, the mark criterion: a process is causal if it is capable of transmitting a local modification in structure, a "mark" (Reichenbach 1956, §23, pp. 197–205; Salmon 1984, 147). The ability to transmit a mark has an asymmetric ring to it; and Salmon says that transmission is a kind of motion (Salmon 1998, 21), which also seems clearly asymmetric. But again, everything depends on the further theoretical transformation of the idea. And here Salmon, faithful to the Humean spirit, avoids introduction of any "mysterious powers". The transmission of a mark from point A in a causal process to point B in the same process *is* the fact that it appears at each point between A and B without further interactions (Salmon 1984, 148). If, as

Salmon still suggests, the transmitted quantities have to remain unchanged through the process (Salmon 1998, 259), causal asymmetry is clearly in danger. That a process has a certain quantity at every stage between A and B cannot be an asymmetric characteristic of that process. Reichenbach correctly observed that when we wish to establish an order among the events  $A_1 \dots A_n$ , it is not sufficient to make a mark in  $A_1$ . “This mark may show in all the events  $A_2 \dots A_n$ ; but we do not thus know the order among these events, and we cannot even tell whether the events are arranged in one causal line”(Reichenbach 1956, 200). So regardless of whether the transmission is manifest or the process only has the ability of transmission, with this at-at view of transmission causal asymmetry does not exist in these models. As it stands this popular approach to causation cannot contribute to our understanding of explanatory asymmetry.

#### 7. MECHANISMS AS ‘COMPLEX’ CAUSE&EFFECT RELATIONS

Under the heading “Mystery ‘mechanism’: an answer too many psychologists like” (Glymour 1998, 41), Clark Glymour claims that many psychologists now think that the answer to the old question of how we learn causes from associations is that we don’t. The reason alluded to is that causes have to do with mechanisms, not with associations. Glymour observes that among psychologists the concept of ‘mechanism’ is seldom further explicated, but from the chosen examples the following picture emerges:

[T]o specify a ‘mechanism’ for a covariation is simply to specify either a sequence of causes that intervene between the candidate cause and effect, where the causal connection posited in the ‘mechanism’ are of a kind that are already familiar and acknowledged. (Glymour 1998, 41)

This view of mechanisms is a sort of cause&effect conception of mechanisms. Mechanisms would perhaps be causal lines with a number of acknowledged intermediary causes and effects. As the following quotation from a recent book

on '*Social mechanisms*' reveals, a similar conception is promoted in sociology:

The search for mechanisms means that we are not satisfied with merely establishing systematic covariation between variables or events; a satisfactory explanation requires that we are also able to specify the social "cogs and wheels" (Elster 1989:3) that have brought the relationship into existence. (Hedström and Swedberg 1998, 7)

Unfortunately, the concept of mechanism that Hedström and Swedberg advocate in their book cannot be trusted in this context. For instance, in chapter two Thomas Schelling claims that "a social mechanism is a plausible hypothesis [...] that could be the explanation of some social phenomenon" (Schelling 1998, pp. 32–33) which simply marries the concept of mechanism to that of explanation. It makes the student of explanation wonder whether anything is gained by turning to this causal approach. The concept of social mechanism seems parasitic on the concept of explanation.

For explanatory purposes a general problem for both these views on causal mechanisms is that they rely on a more fundamental conception of causation; they are content to substitute a more complex picture of the underlying causes and effects for the regularities that occur on the surface level, and they typically understand the underlying causes and effects according to one of the traditional views we have discussed above. For this reason these mechanistic views belong to a theoretically redundant category.<sup>9</sup> If the more fundamental conception provides asymmetry then these psychological and social mechanisms will be able to take over and elaborate on this idea; but if the conceptions they build on lack asymmetry, then nothing in this concept of mechanism will add it. Also this causal approach fails to contribute asymmetry to explanations.

#### 8. A LAST RESORT. MECHANISMS AS DISPOSITIONS OR CAPACITIES

To repeat, the aim of this paper has been to show how traditional and current

theoretical transformations of the causal concept make it a serious mistake to reason as if its “pre-theoretic” connotations are preserved in a way that for instance makes causal models of explanation contribute asymmetry. All of the above causal approaches seem to have a problem in preserving an adequate kind of asymmetry.

I might have been too quick in disregarding some of these approaches—perhaps the right kind of fixes can be developed or already exist. However, the number of illustrations is compelling evidence for the claim that one runs a serious risk of being misled in this context. There is the risk of accepting a causal model for characteristics it doesn’t have; and there is the risk of thinking that it does better than its predecessors.

This doesn’t mean that I think every causal theory has to fail as a model for explanatory asymmetry. I will therefore end this paper by an equally brief presentation of another sense of ‘mechanism’ that should be contemplated by the student of explanation. ‘True’ mechanisms would not be ready-made processes but—using Nancy Cartwright’s characterisation—rather the machines that can be realised as such processes:

What is a nomological machine? It is a fixed (enough) arrangement of stable (enough) capacities that in the right sort of stable (enough) environment will, with repeated operation, give rise to the kind of regular behaviour that we represent in our scientific laws. (Cartwright 1999, 28)

A causal-mechanism view in this second sense makes concepts such as ‘capacities’ (Cartwright 1999), ‘dispositions’ (as employed in Martin 1994), or ‘causability’ (Mellor 1995) fundamental. A causal-mechanism view in Glymour’s, Hedström’s and Swedberg’s sense still takes ‘causes’ to be the fundamental concept; this is where they differ. To uphold the difference one needs to refrain from accounting for capacities, dispositions, and causability in terms of causal or other relations. That it is possible is plain: One can always deny that they are reducible to their manifestations.<sup>10</sup> For the sake of illustration I stipulate the irreducibility of mechanisms. What interests me then

are two things:

First, the move should not be ad hoc. Since there are plenty of reasons why we need irreducible mechanisms this should not worry us. Causal situations where either nothing happens or the cause or effect is “negative” are independent and even more important reasons (Mellor 1995). Especially for explanatory purposes, an account that admitted negative causation would be of considerable value.

Second, we need some preliminary reason why a causal-mechanism view in the second sense furthers our understanding of causal asymmetry where other approaches fail. As we have seen, many of the other candidates have a problem with asymmetry because what was pre-theoretically assumed to be a local intrinsic affair, turned out to be non-local and extrinsic. Since a mechanism in the second sense is not reducible to its manifestation and ought to be located where the local causal situation we are interested in exists, it could be both local and intrinsic. Thus it might have the two characteristics we have complained were lost in other approaches to causation. This does however not prove what is most important, namely that it can account for asymmetry in this local intrinsic way as well. However, there seems to be many ways an appropriate kind of asymmetry could be had once mechanisms in the second sense are accepted.

Let us first assume that causal mechanisms, though not reducible to their manifestations, only works once (and are then destroyed). Now, if our pre-theoretic idea about causes explaining their effects but not *vice versa* is adequate, then nothing can possibly have been transformed by moving to this conception of causation as a single-case mechanism. Why did the firework go off? Because I lit the fuse. We have the cause, we have the effect, and since we have a causal mechanism that only worked once, the reason why the effect occurred was that the cause occurred. Such once-and-for-all mechanisms may be objected to on other grounds. There are epistemological constraints that should be considered. Can we know that there are such mechanisms and that they work in one way rather than the other? Well, since we examined the

firework in advance we knew that the mechanism wasn't a fake; I felt the pain when my thumb came in the way of the burning fuse. We followed the rest of the process by eyesight, etc. There seems to be a lot of positive knowledge about this firework and how it worked. Sometimes we also experience causal forces directly, as when we try to swim in a current (Fales 1990, Persson 1997). The epistemology of causation is in many situations no more mysterious if we assume single-case mechanisms than if we stick to one of the many other approaches we have discussed in this paper. Yet asymmetry is not transformed beyond recognition.

But wouldn't this kind of asymmetry have as high a price as the asymmetric manipulation laws we discussed in connection with Gasking? Dowe sketches a "fundamental" objection to any intrinsic account of causal asymmetry.

[Such a view] would entail that the current laws of physics are radically incomplete, for this new physical theory would identify laws of nature which are in some sense not time symmetric. (Dowe 1996, 246)

The problem with Dowe's objection is that the physicist's symmetry and asymmetry he mentions occurs on the law-level, while causal mechanisms in the second sense operate on the local level—and so does as Lipton remarked above often explanations. The objection could therefore be immediately relevant only to someone who accounted for causal asymmetry at the local level by referring to the law-level (Gasking might be a good example). To apply to mechanisms in the second sense it would presuppose a conception similar to Armstrong's view of causation as a second-order relation between (universals in) first-order facts (Armstrong 1997).

Dowe's objection can be reformulated, though. If causal mechanisms in the second sense generate physical laws, and that, as Dowe's position implies, these laws are symmetrical, then how can these mechanisms be causally asymmetric? Does that which occurs on the local level conflict with what goes

on in the special sciences? If this is what Dowe's objection boils down to, it is something well-known and not very alarming. We often have symmetric correlations although we are convinced that they occur because of asymmetric causal relations. This is the very point of adopting mechanistic views in psychology and sociology. It is first when we make the further assumption that the causal mechanisms or relations are to be accounted for in the same sort of law-like regularities our correlations were expressed in that a problem emerges. By keeping the levels apart, as our present concept urges us to, it is obvious why there need be no conflict between local single-case mechanisms and general symmetry.

But we should not be content with causal asymmetry in worlds where we only have single-case mechanisms. Mechanisms may come in a varied assortment (Persson 1999). Would mechanisms that work several times and where we can use the cause to produce the effect as easily as the other way around be more problematic? Let us assume that it is one and the same psychological mechanism in me that responds to failure by causing anxiety and to anxiety by causing failure.

The simplest case would be to have such a "symmetric" mechanism that loses its function when having performed. Even though the mechanism could have worked in either of the two ways, this does not mean it did so in the single case. Then the possibility of working either way seems to be fully sufficient for the general-symmetry claim while what happened in the individual case was clearly asymmetric.

This case could be generalised since what is needed to preserve asymmetry in every particular case is only that even if P is causable by Q and Q is causable by P, the realisation of one of these possible causal relations always precludes the existence of the other (Mellor 1995, 225–229) Whether or not such a result can be proved depends on the specific theory of causal mechanism in the second sense, and this is not the right place to go into more detail. And I don't need to because I am convinced that *for explanatory purposes* a much weaker result is required. Let us assume that a particular

causal loop is possible, i.e. that we have a true case where A causes B and B causes A, and furthermore that the reason is that a certain causal mechanism works this way. It is indeed a symmetric causal dependence on the local level. Well, in that case I would claim that we simply don't have explanatory asymmetry with regard to causes and effects. What we might have is asymmetry regarding causal mechanism and effect, but that is another issue.

#### CONCLUSION

Contrary to what many of us might think, the reason why there is explanatory asymmetry is not causal asymmetry. At least not if causation is what some of the most popular theories of causation say it is. As was hinted at in the introduction, locality may behave similarly so this result is an instance of a more general problem. What we can learn from this is that theories of causation are rather tested than applied in cases like these, where we rely on causal models in accounting for this or that. Even if the present evidence is not robust enough for justifying the conclusion that what we presuppose when following Lipton in thinking that causal models account for explanatory asymmetry is that our concept of causation is linked to mechanisms in the second sense, that picture is at least suggested. Whether such a view is acceptable depends on its other consequences. Similarly, although many of the above-mentioned theories do not pass this test, they might of course have other virtues that on the whole make them preferable. But then students of explanation should turn elsewhere for an account of explanatory asymmetry.

## LITERATURE

- Armstrong, D. M. (1997), *A world of states of affairs*. Cambridge: Cambridge University Press.
- Bird, Alexander (1999), "Explanation and laws", *Synthese* **120**(1): 1–18.
- Cartwright, Nancy (1999), *The dappled world*. Cambridge: Cambridge University Press.
- Dowe, Phil (1992), "Wesley Salmon's process theory of causality and the conserved quantity theory", *Philosophy of science* **62**: 195–216.
- Dowe, Phil (1996), "Backwards causation and the direction of causal processes", *Mind* **105**(418): 227–248.
- Elster, Jon (1989), *Nuts and bolts for the social sciences*. Cambridge: Cambridge University Press.
- Fales, Evan (1990), *Causation and universals*. London: Routledge.
- Gasking, Douglas (1955), "Causation and recipes", *Mind* **64**: 479–87.
- Glymour, Clark (1998), "Learning causes: Psychological explanations of causal explanation", *Minds and machines* **8**: 39–60.
- Hausman, Daniel M. (1998), *Causal asymmetries*. Cambridge: Cambridge University Press.
- Hedström, Peter and Richard Swedberg, (eds.) (1998), *Social mechanisms*. Cambridge: Cambridge University Press.
- Hempel, Carl G. (1965), "Aspects of scientific explanation", in his *Aspects of scientific explanation*. New York: The Free Press.
- Hempel, Carl G. and Paul Oppenheim (1948/1965), "Studies in the logic of explanation", in C. G. Hempel, *Aspects of scientific explanation*. New York: The Free Press.
- Hume, David (1777/1902), *Enquiries concerning the human understanding and concerning the principles of morals*. Oxford: Clarendon Press.
- Lewis, David (1979/1986), "Counterfactual dependence and time's arrow", in his *Philosophical papers*. Oxford: Oxford University Press, **II**: 32–66.

- Lipton, Peter (1991), *Inference to the best explanation*. London: Routledge.
- Martin, C. B. (1994), "Dispositions and conditionals", *The philosophical quarterly* **44**: 1–8.
- Mellor, D. Hugh (1995), *The facts of causation*. London: Routledge.
- Persson, Johannes (1997), *Causal facts*. Stockholm: Thales.
- Persson, Johannes (1999), "The determinables of explanatory mechanisms", *Synthese* **120**(1): 77–87.
- Price, Huw (1992), "Agency and causal asymmetry", *Mind* **101**(403): 501–520.
- Price, Huw (1996), "Backwards causation and the direction of causal processes: Reply to Dowe", *Mind* **405**(419): 467–475.
- Reichenbach, Hans (1956), *The direction of time*. Berkeley and Los Angeles: University of California Press.
- Russell, Bertrand (1912/1954), "On the notion of cause", in his *Mysticism and logic*. London: George Allen & Unwin Ltd: 180-208.
- Russell, Bertrand (1948/1951), *Human knowledge*. London: George Allen and Unwin Ltd.
- Salmon, Wesley (1984), *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Salmon, Wesley, C. (1998), *Causality and explanation*. New York: Oxford University Press.
- von Wright, G. H. (1975), *Causality and determinism*. New York: Columbia University Press.

## NOTES

---

<sup>1</sup> I have borrowed this particular example from Bird 1999, and variations on the theme can be found in numerous discussions concerning the problems with the D-N model.

<sup>2</sup> There does not seem to be any unity in the kinds of "things" we tend to explain. And it is not always good methodology to increase unity. Just as we have kinetic and potential energy in Newtonian physics, advances in the study of explanation might be dependent on the distinction between different kinds of explanation.

<sup>3</sup>Hempel and Oppenheim 1948/1965, 250 and Hempel 1965, 347–354.

---

<sup>4</sup> Unless it is projected onto them, and then the difference between von Wright, Gasking, and Price might vanish.

<sup>5</sup> This has been noted by many. See Price 1992, Dowe 1996, and Hausman 1998.

<sup>6</sup> "In the first place, the region to which we infer need not be later than those from which we infer" (Russell 1948/1951, 326); similar passages can be found in Russell 1912/1954.

<sup>7</sup> It has its forerunners too in the writings of Quine and Fair, for instance, but that is of more importance for those interested in exegesis of ideas.

<sup>8</sup> Another problem is, as Alexander Bird made me aware of, that there may be some 'grue'-like quantity that is conserved as a matter of accidental regularity.

<sup>9</sup> By the way: their close relation to the before mentioned categories will almost immediately prove those psychologists wrong. Repeated cause&effect relations will certainly give rise to some associations or patterns, and associations will therefore also be valuable as important evidence for such mechanisms, "a disconnection between mechanisms, on the one hand, and probabilistic patterns, on the other, puts everything on a false footing" (Glymour 1998, 43).

<sup>10</sup> "What makes capacity claims true are claims about capacities [...]" (Cartwright 1999, 72).